Projection-Based Wavefunction-in-DFT Embedding

Sebastian J. R. Lee,[†] Matthew Welborn,[†] Frederick R. Manby,[‡] and Thomas F. Miller III[*,†]

† *Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, CA 91125*

‡ *Centre for Computational Chemistry, School of Chemistry, University of Bristol, Bristol BS8 1TS, United Kingdom*

E-mail: tfm@caltech.edu

Conspectus

Complex chemical systems present challenges to electronic structure theory, stemming from large system sizes, subtle interactions, coupled dynamical timescales, or electronically non-adiabatic effects. New methods are needed to perform reliable, rigorous, and affordable electronic structure calculations for simulating the properties and dynamics of such systems.

This Account reviews projection-based quantum embedding for electronic structure. The method provides a simple, robust, and accurate approach for describing a small part of a chemical system at the level of a correlated wavefunction (WF) method while the remainder of the system is described at the level of density functional theory (DFT). We present the theoretical underpinnings of projection-based embedding, describe use of the method for combining wavefunction and density-functional theories, and discuss technical refinements that have improved the applicability and robustness of the method. Applications of projection-based WF-in-DFT embedding are also reviewed, with particular focus on recent work on transition-metal catalysis, enzyme reactivity, and battery electrolyte decomposition. Looking forward, we anticipate continued refinement of the projection-based embedding methodology, as well as increasingly widespread application in diverse areas of chemistry, biology, and materials science.

Introduction

Computational modeling of chemical and photochemical processes in complex systems faces extraordinary challenges from the perspective of electronic structure theory. Target applications span frontier areas in chemistry and materials science, including photosynthesis,[1] biocatalysis,[2] electrochemistry,[3] geochemistry,[4] and surface science.[5] These problems combine large system sizes with subtle interactions, and in many cases, they involve multiple dynamical timescales and electronically non-adiabatic effects. The development of new methods to perform reliable, rigorous, and affordable electronic structure calculations for the simulation of dynamics in large systems remains the central challenge in theoretical chemistry.

The use of density functional theory (DFT) for the description of electronic structure has gained remarkable prevalence in recent years, due to its reasonable compromise between ac-

curacy and computational cost.[6] However, the chemical sciences are permeated with systems for which the approximations of DFT fundamentally break down or for which the computational cost of DFT remains prohibitive for the MD simulation of necessary length- and timescales. Although correlated wavefunction (WF) electronic structure methods, such as coupled-cluster theory, provide better accuracy than DFT for single point calculations on systems of modest size, they have been too expensive to allow for widespread use in terms of exploring conformational landscapes and reaction pathways.

To mitigate the trade-off between accuracy and computational cost, quantum embedding has emerged as a powerful strategy for modeling the electronic structure of complex systems. In embedding methods, a high-level quantum-mechanical description of a chemically active subsystem is embedded in a surrounding environment described using a more approximate theory. By exploiting the intrinsic locality of molecular interactions, this approach provides high accuracy for regions that demand it while avoiding the computational cost of a high-level calculation on the whole system. Notable examples of embedding include QM/MM,[7,8] ONIOM,[9] fragmentation methods,[10–12] density functional embedding,[13–29] and density matrix embedding,[30–33] although there are many manifestations of the idea.

Projection-based embedding[27] describes subsystem interactions at the level of DFT and allows for the partitioning of the subsystems across covalent and even conjugated bonds, and it enables the use of relatively small subsystem sizes for an embedded WF description. The current Account aims to provide a practical review of the projection-based embedding method, including a description of its theory, implementation, applications, and limitations. Although we describe the methodological context for projection-based embedding, we also direct the reader to several reviews that provide a more complete description of alternative approaches.[9,10,34–38]

Projection-Based Embedding

Quantum embedding methods developed within the framework of DFT offer a formally exact approach to electronic structure calculations in which complex chemical problems are decomposed into the solution of individual smaller subsystems.[19,38] Throughout this review, we shall use the term "exact" to denote that a DFT-in-DFT embedding calculation where both subsystems are treated using the same exchange-correlation (XC) functional yields the same result as a single Kohn-Sham (KS) DFT calculation performed over the full system. In principle, DFT embedding thus avoids the uncontrolled approximations (such as link atoms) that appear in widely used methods, such as QM/MM and ONIOM.

In practice, however, many DFT embedding studies employ substantial approximations in the description of subsystem interactions. The subsystem interaction potentials that emerge in the DFT embedding framework include non-additive kinetic potential (NAKP) terms that enforce Pauli exclusion between the electrons of the various subsystems.[38] Without knowledge of the exact functional for the non-interacting kinetic energy, this has typically required approximate NAKP treatments that break down in cases for which the subsystem densities significantly overlap (which include hydrogen-bonded or covalently bonded subsystems),[38–40] limiting applications to those involving weakly interacting subsystems. Although numerically exact DFT embedding methods have been developed that determine NAKP contributions via an optimized effective potential (OEP) inversion of the density,[18,21,24,25,41–43] OEP inversion can be ill-conditioned and requires careful regularization protocols.[18,26,44–47]

Projection-based embedding avoids these issues by providing a numerically exact DFT-in-DFT embedding framework that eliminates the NAKP contributions via the mutual orthogonalization of the subsystem molecular orbitals. Panel (a) of Fig. 1 outlines the general procedure of a projection-based embedding calcula-

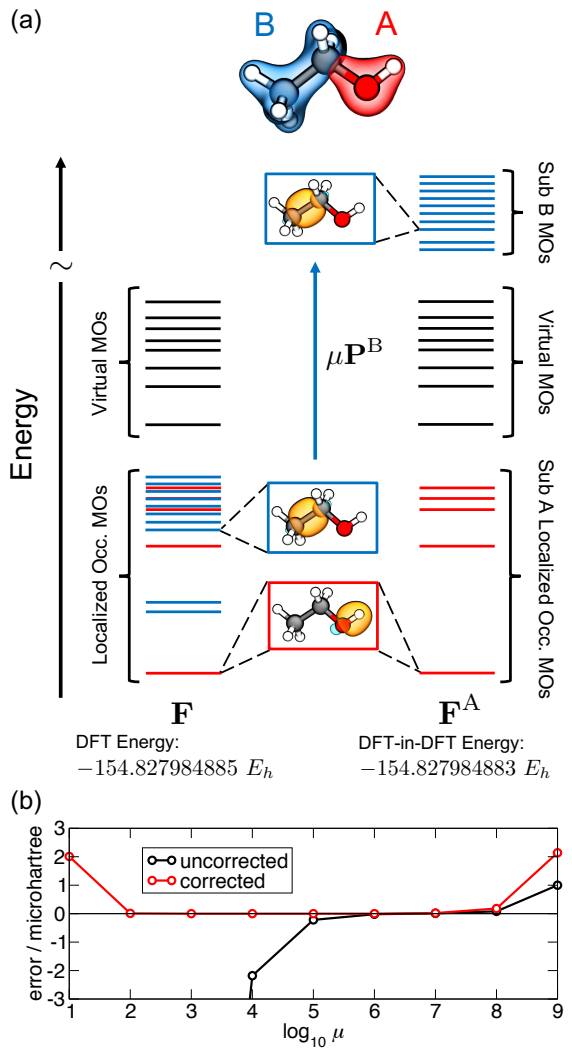

Figure 1: (a) Demonstration of projection-based embedding, using the example of embedding the 10 electrons of the –OH moiety of ethanol in the environment of the ethyl subsystem. $\mathbf{F}$ is the full system KS-DFT Fock matrix that is initially used to self-consistently determine the occupied KS MOs of the full system. The occupied MOs are then localized and grouped into subsystems A and B (red and blue respectively). $\mathbf{F}^{\mathrm{A}}$ is the embedded Fock matrix (Eq. 4) for the subsystem A electrons, which includes the projection operator, $\mu\mathbf{P}^{\mathrm{B}}$ (Eq. 5). Following projection, the subsystem A LMOs (red) are explicitly orthogonalized with respect to the subsystem B LMOs (blue), thus eliminating non-additive kinetic energy contributions. (b) Error in the uncorrected (Eq. 1) and corrected (Eq. 1 + Eq. 8) PBE-in-PBE/6-31G* energy expressions relative to full KS-DFT on ethanol using PBE/6-31G*, demonstrating that the perturbative correction (Eq. 8) yields essentially exact embedding energies over a wide range of μ values. Adapted with permission from Ref. 27. Copyright 2012 American Chemical Society.

tion. A KS-DFT calculation is first performed on the full system to self-consistently determine the KS orbitals and the corresponding Fock matrix, $\mathbf{F}$. The occupied KS orbitals are then localized, shown on the left side of panel (a) of Fig. 1. These localized molecular orbitals (LMOs) are partitioned into subsystems A (in red) and B (in blue), with corresponding atomic-orbital-basis density matrices, $\boldsymbol{\gamma}^{\mathrm{A}}$ and $\boldsymbol{\gamma}^{\mathrm{B}}$.

To determine the subsystem Fock matrix, $\mathbf{F}^{\mathrm{A}}$, that describes the electrons of subsystem A in the environment of the density matrix of subsystem B, we begin with the DFT-in-DFT energy expression for projection-based embedding,

$$\begin{aligned} E_{\text{DFT-in-DFT}}\left[\tilde{\boldsymbol{\gamma}}^{\mathrm{A}};\boldsymbol{\gamma}^{\mathrm{A}},\boldsymbol{\gamma}^{\mathrm{B}}\right] &= E_{\text{DFT}}\left[\tilde{\boldsymbol{\gamma}}^{\mathrm{A}}\right] \\ &+ E_{\text{DFT}}\left[\boldsymbol{\gamma}^{\mathrm{A}}+\boldsymbol{\gamma}^{\mathrm{B}}\right] - E_{\text{DFT}}\left[\boldsymbol{\gamma}^{\mathrm{A}}\right] \\ &+ \text{tr}\left[(\tilde{\boldsymbol{\gamma}}^{\mathrm{A}}-\boldsymbol{\gamma}^{\mathrm{A}})\mathbf{v}_{\text{emb}}\left[\boldsymbol{\gamma}^{\mathrm{A}},\boldsymbol{\gamma}^{\mathrm{B}}\right]\right] \\ &+ \mu\text{tr}\left[\tilde{\boldsymbol{\gamma}}^{\mathrm{A}}\mathbf{P}^{\mathrm{B}}\right], \end{aligned} \tag{1}$$

where E_{DFT} denotes the KS-DFT energy evaluated using the bracketed density matrix, $\tilde{\boldsymbol{\gamma}}^{\mathrm{A}}$ is the embedded subsystem A density matrix, and $\mathbf{P}^{\mathrm{B}}$ is a projection operator that enforces the mutual orthogonalization of subsystem A and B LMOs (see below). The embedding potential, $\mathbf{v}_{\text{emb}}$, describes all interactions between subsystems A and B,

$$\mathbf{v}_{\text{emb}}\left[\boldsymbol{\gamma}^{\mathrm{A}},\boldsymbol{\gamma}^{\mathrm{B}}\right] = \mathbf{g}\left[\boldsymbol{\gamma}^{\mathrm{A}}+\boldsymbol{\gamma}^{\mathrm{B}}\right] - \mathbf{g}\left[\boldsymbol{\gamma}^{\mathrm{A}}\right]. \tag{2}$$

In general, $\mathbf{v}_{\text{emb}}$ would also include the difficult-to-evaluate NAKP contributions, but if the subsystem densities are constructed from disjoint subsets of orthogonal orbitals, these NAKP terms are exactly zero.[27] The matrix $\mathbf{g}$ is the density-matrix functional of two-electron terms, given by

$$\begin{aligned} \left(\mathbf{g}\left[\boldsymbol{\gamma}\right]\right)_{\kappa\nu} &= \sum_{\lambda\sigma}\gamma_{\lambda\sigma}\left[(\kappa\nu|\lambda\sigma) - \tfrac{1}{2}x(\kappa\lambda|\nu\sigma)\right] \\ &+ \left(\mathbf{v}_{\text{xc}}\left[\boldsymbol{\gamma}\right]\right)_{\kappa\nu}, \end{aligned} \tag{3}$$

where κ, ν, λ and σ label atomic orbital basis functions, $(\kappa\nu|\lambda\sigma)$ are two-electron repulsion integrals, x is the fraction of exact exchange,

and $\mathbf{v}_{\text{xc}}$ is the exchange-correlation potential matrix.

The subsystem Fock matrix corresponding to variation of equation 1 with respect to $\tilde{\boldsymbol{\gamma}}^{\text{A}}$ is then

$$\begin{aligned}\mathbf{F}^{\text{A}} &= \frac{\partial}{\partial \tilde{\boldsymbol{\gamma}}^{\text{A}}} E_{\text{DFT-in-DFT}}\left[\tilde{\boldsymbol{\gamma}}^{\text{A}}; \boldsymbol{\gamma}^{\text{A}}, \boldsymbol{\gamma}^{\text{B}}\right] \\ &= \mathbf{h} + \mathbf{g}\left[\tilde{\boldsymbol{\gamma}}^{\text{A}}\right] + \mathbf{v}_{\text{emb}}\left[\boldsymbol{\gamma}^{\text{A}}, \boldsymbol{\gamma}^{\text{B}}\right] + \mu \mathbf{P}^{\text{B}},\end{aligned} \tag{4}$$

where $\mathbf{h}$ is the standard one-electron Hamiltonian. Self-consistent optimization of $\mathbf{F}^{\text{A}}$ with respect to $\tilde{\boldsymbol{\gamma}}^{\text{A}}$ recovers the original subsystem A density matrix, $\boldsymbol{\gamma}^{\text{A}}$ (Fig. 1, right column) for the case of DFT-in-DFT embedding when both subsystems are described using the same XC functional.

A practical way to enforce the orthogonality of the subsystem A orbitals to those in subsystem B is to introduce a level-shift operator of the form[27]

$$\mu \mathbf{P}^{\text{B}} = \mu \mathbf{S} \boldsymbol{\gamma}^{\text{B}} \mathbf{S}, \tag{5}$$

where $\mathbf{S}$ is the overlap matrix in the atomic orbital basis, and μ is a positive scalar number. The action of this operator is to level-shift the subsystem B LMOs to high energies so that they cannot hybridize with those of subsystem A (shown on the right side of panel (a) of Fig. 1).[48,49] In the $\mu \rightarrow \infty$ limit, Eq. 1 reduces to the KS-DFT energy for the full system, such that the projection-based approach is exact for DFT-in-DFT embedding.

Embedding methods that maintain orthogonality between subsystem orbitals have long been in use, including the Philips-Kleinman pseudopotential approach.[48] What had not been previously recognized is that these same strategies can be used to formulate a formally exact method for DFT embedding.[27]

DFT-in-DFT Embedding

Before proceeding, it is worth expanding on several aspects of DFT-in-DFT embedding. First, it is clear from the preceding discussion that the projection-based approach allows for a description of DFT-in-DFT embedding with subsystems A and B evaluated using different XC functionals. Typically this involves using a more expensive (i.e. hybrid, meta-GGA, etc.) functional to describe subsystem A and a computationally cheaper (i.e. GGA or LDA) functional to describe subsystem B.

The procedure for this type of DFT-in-DFT embedding calculations begins with performing a low-level KS-DFT calculation on the full system, yielding $E_{\text{DFT}}[\boldsymbol{\gamma}^{\text{A}} + \boldsymbol{\gamma}^{\text{B}}]$. The resulting occupied MOs are localized and partitioned into subsystems A and B, which are used to form the matrices $\boldsymbol{\gamma}^{\text{A}}$, $\boldsymbol{\gamma}^{\text{B}}$, $\mathbf{P}^{\text{B}}$, and $\mathbf{v}_{\text{emb}}[\boldsymbol{\gamma}^{\text{A}}, \boldsymbol{\gamma}^{\text{B}}]$ and to evaluate $E_{\text{DFT}}[\boldsymbol{\gamma}^{\text{A}}]$; these quantities are unchanged during the self-consistent field (SCF) iterations for the embedded subsystem. The SCF iterations for the embedded subsystem are performed to optimize the subsystem density matrix, $\tilde{\boldsymbol{\gamma}}^{\text{A}}$. At each SCF iteration, $\mathbf{F}^{\text{A}}$ is calculated (Eq. 4) and diagonalized; $\mathbf{g}[\tilde{\boldsymbol{\gamma}}^{\text{A}}]$ is the only term in Eq. 4 to be re-evaluated at each SCF iteration, and it is done so using the high-level XC functional. Finally, to obtain the total DFT-in-DFT energy, the converged subsystem density, $\tilde{\boldsymbol{\gamma}}^{\text{A}}$, is used to evaluate $E_{\text{DFT}}[\tilde{\boldsymbol{\gamma}}^{\text{A}}]$ using the high-level XC functional, as well as the last two traces on the right hand side of Eq. 1.

In light of the quantities that must be iteratively re-evaluated during an embedding calculation that involves two different levels of theory, we provide a few additional comments regarding the derivation of the total DFT-in-DFT energy expression. The starting point for Eq. 1 is the more transparent energy expression

$$\begin{aligned}&E_{\text{DFT-in-DFT}}\left[\tilde{\boldsymbol{\gamma}}^{\text{A}}; \boldsymbol{\gamma}^{\text{B}}\right] = \\ &\quad E_{\text{DFT}}\left[\tilde{\boldsymbol{\gamma}}^{\text{A}} + \boldsymbol{\gamma}^{\text{B}}\right] + \mu \text{tr}\left[\tilde{\boldsymbol{\gamma}}^{\text{A}} \mathbf{P}^{\text{B}}\right].\end{aligned} \tag{6}$$

Minimization of this total energy expression with respect to the density matrix for subsystem A would lead to an expression for the subsystem Fock matrix $\mathbf{F}^{\text{A}}$ that involves the costly re-evaluation of the embedding potential in terms of the high-level XC functional at each SCF iteration. To avoid this, a first-order expansion in $\tilde{\boldsymbol{\gamma}}^{\text{A}} - \boldsymbol{\gamma}^{\text{A}}$ is performed, yielding

$$\begin{aligned}&E_{\text{DFT}}\left[\tilde{\boldsymbol{\gamma}}^{\text{A}} + \boldsymbol{\gamma}^{\text{B}}\right] - E_{\text{DFT}}\left[\tilde{\boldsymbol{\gamma}}^{\text{A}}\right] \approx \\ &\quad E_{\text{DFT}}\left[\boldsymbol{\gamma}^{\text{A}} + \boldsymbol{\gamma}^{\text{B}}\right] - E_{\text{DFT}}\left[\boldsymbol{\gamma}^{\text{A}}\right] \\ &\quad + \text{tr}\left[(\tilde{\boldsymbol{\gamma}}^{\text{A}} - \boldsymbol{\gamma}^{\text{A}}) \mathbf{v}_{\text{emb}}\left[\boldsymbol{\gamma}^{\text{A}}, \boldsymbol{\gamma}^{\text{B}}\right]\right].\end{aligned} \tag{7}$$

Rearranging Eq. 7 and substituting it into Eq. 6 yields Eq. 1. Note that if both subsystems are described using the same XC functional, this perturbative approximation becomes exact in the limit of mutual orthogonalization of the subsystem orbitals, since $\tilde{\boldsymbol{\gamma}}^{\mathrm{A}}$ approaches $\boldsymbol{\gamma}^{\mathrm{A}}$. If different exchange-correlation functionals are employed for the two subsystems, then Eq. 7 can lead to density-driven errors associated with the fact that the density matrix obtained using the low-level theory is different from that obtained using the high-level theory.[20,26,50,51]

Finally, we discuss the convergence of the projection-based embedding description with respect to the level-shift parameter, μ. Although the projection operator in Eq. 4 only exactly enforces orthogonality between subsystem A and B orbitals in the limit of $\mu \to \infty$, finite values of μ in the range of 10^4 a.u. to 10^7 a.u. are consistently found to yield accurate results (panel (b) Fig. 1, black), regardless of chemical system;[27] a default choice of $\mu = 10^6$ a.u. has been found in almost all cases to yield microhartree-scale embedding errors. Nonetheless, if greater accuracy is needed, then a perturbative correction

$$\mu \mathrm{tr}\left[\tilde{\boldsymbol{\gamma}}^{\mathrm{A}}\mathbf{P}^{\mathrm{B}}\right] \tag{8}$$

can be added to the DFT-in-DFT energy expression to account for the finiteness of μ,[27] and this typically leads to sub-microhartree accuracy over a very large range of μ values (panel (b) Fig. 1, red). Errors associated with finite values of μ can also be avoided by enforcing the projection via explicit orthogonalization[52–56] of the subsystem orbitals, at some cost to the simplicity of the implementation.

Wavefunction-in-DFT Embedding

Beyond DFT-in-DFT embedding, the projection-based approach readily allows for wavefunction-in-DFT (WF-in-DFT) embedding, in which subsystem A is treated using a WF-level description and subsystem B is described at the DFT level.[27] Starting from the DFT-in-DFT energy expression in Eq. 1, the WF-in-DFT energy is simply obtained by substituting the DFT energy of subsystem A with the corresponding WF energy,

$$\begin{aligned} E_{\text{WF-in-DFT}}\left[\tilde{\Psi}^{\mathrm{A}};\boldsymbol{\gamma}^{\mathrm{A}},\boldsymbol{\gamma}^{\mathrm{B}}\right] &= E_{\mathrm{WF}}\left[\tilde{\Psi}^{\mathrm{A}}\right] \\ &+ \mathrm{tr}\left[(\tilde{\boldsymbol{\gamma}}^{\mathrm{A}}-\boldsymbol{\gamma}^{\mathrm{A}})\mathbf{v}_{\mathrm{emb}}\left[\boldsymbol{\gamma}^{\mathrm{A}},\boldsymbol{\gamma}^{\mathrm{B}}\right]\right] \\ &+ E_{\mathrm{DFT}}\left[\boldsymbol{\gamma}^{\mathrm{A}}+\boldsymbol{\gamma}^{\mathrm{B}}\right] - E_{\mathrm{DFT}}\left[\boldsymbol{\gamma}^{\mathrm{A}}\right] \\ &+ \mu\mathrm{tr}\left[\tilde{\boldsymbol{\gamma}}^{\mathrm{A}}\mathbf{P}^{\mathrm{B}}\right], \end{aligned} \tag{9}$$

where $\tilde{\Psi}^{\mathrm{A}}$ is the WF for subsystem A, $\tilde{\boldsymbol{\gamma}}^{\mathrm{A}}$ is the one-particle reduced density matrix corresponding to $\tilde{\Psi}^{\mathrm{A}}$, and $E_{\mathrm{WF}}[\tilde{\Psi}^{\mathrm{A}}]$ is the WF energy of subsystem A.[27,29]

A projection-based WF-in-DFT embedding calculation proceeds as follows. A KS-DFT calculation is first performed over the full system. The resulting occupied MOs are localized and partitioned into two sets, corresponding to subsystems A and B. These sets are used to construct $\mathbf{h}^{\text{A-in-B}}$,

$$\mathbf{h}^{\text{A-in-B}}\left[\boldsymbol{\gamma}^{\mathrm{A}},\boldsymbol{\gamma}^{\mathrm{B}}\right] = \mathbf{h} + \mathbf{v}_{\mathrm{emb}}\left[\boldsymbol{\gamma}^{\mathrm{A}},\boldsymbol{\gamma}^{\mathrm{B}}\right] + \mu\mathbf{P}^{\mathrm{B}}, \tag{10}$$

which is an effective one-electron Hamiltonian containing the standard one-electron Hamiltonian, the embedding potential and the projection operator. Finally, a correlated WF calculation is performed on subsystem A wherein $\mathbf{h}^{\text{A-in-B}}$ replaces the standard one-electron Hamiltonian. The final WF-in-DFT energy is given by equation 9.

The WF calculation for subsystem A consists of two steps: first, a set of reference orbitals is generated and second, a correlated WF calculation is performed using those orbitals. The reference orbitals can be obtained either via Hartree-Fock (HF) or a multiconfigurational method. For the former case, the subsystem A post-HF calculation begins with HF-in-DFT embedding. The HF-in-DFT Fock matrix, $\mathbf{F}^{\mathrm{A}}$, is derived by inserting a Slater determinant for the subsystem A WF into Eq. 9 and differentiating with respect to $\tilde{\boldsymbol{\gamma}}^{\mathrm{A}}_{\mathrm{HF}}$ giving

$$\begin{aligned} \mathbf{F}^{\mathrm{A}} &= \frac{\partial}{\partial\tilde{\boldsymbol{\gamma}}^{\mathrm{A}}_{\mathrm{HF}}} E_{\text{HF-in-DFT}}\left[\tilde{\boldsymbol{\gamma}}^{\mathrm{A}}_{\mathrm{HF}};\boldsymbol{\gamma}^{\mathrm{A}},\boldsymbol{\gamma}^{\mathrm{B}}\right] \\ &= \mathbf{h}^{\text{A-in-B}}\left[\boldsymbol{\gamma}^{\mathrm{A}},\boldsymbol{\gamma}^{\mathrm{B}}\right] + \mathbf{g}\left[\tilde{\boldsymbol{\gamma}}^{\mathrm{A}}_{\mathrm{HF}}\right], \end{aligned} \tag{11}$$

where $\mathbf{g}$ includes all of the usual HF two-electron terms, and $\mathbf{h}^{\text{A-in-B}}$ represents the effective one-electron Hamiltonian given by Eq. 10. Once the subsystem A HF MOs are optimized in the presence of the DFT embedding potential, $\mathbf{v}_{\text{emb}}$, they are used for the correlated subsystem A post-HF calculation. An analogous procedure holds for the case of multireference methods, wherein a multiconfigurational WF is substituted in place of the single Slater determinant in Eq. 9.[57,58] In this way, projection-based WF-in-DFT embedding can be readily performed with any existing WF method (or quantum impurity solver) simply by modifying the one-electron Hamiltonian in the WF method to include the projection-based embedding terms.

While projection-based embedding is exact for (same functional) DFT-in-DFT embedding, projection-based WF-in-DFT embedding is necessarily approximate. Some sources of error in WF-in-DFT embedding have been analyzed, including the approximate nature of the non-additive exchange-correlation energy[29] and density-driven errors in the underlying DFT calculation.[51]

Analytical Nuclear Gradients

Analytical nuclear gradients for projection-based WF-in-DFT embedding have been developed[59] and implemented in the Molpro quantum chemistry package,[60] enabling exploration of the PES, including geometry optimization and minimum energy reaction pathway searches.

The gradient theory is formulated in terms of a Lagrangian

$$\begin{aligned}\mathcal{L}\left[\mathbf{C},\tilde{\Psi}^{\text{A}},\mathbf{\Lambda},\mathbf{x},\mathbf{z}^{\text{loc}},\mathbf{z}\right] &= \\ E_{\text{WF-in-DFT}}\left[\tilde{\Psi}^{\text{A}};\boldsymbol{\gamma}^{\text{A}},\boldsymbol{\gamma}^{\text{B}}\right] &+ \sum_{m}\Lambda_{m}^{\text{WF,A}}c_{m} \\ + \sum_{pq}x_{pq}[\mathbf{C}^{\dagger}\mathbf{S}\mathbf{C}-\mathbf{1}]_{pq} &+ \sum_{ai}z_{ai}F_{ai} + \sum_{i>j}z_{ij}^{\text{loc}}r_{ij}\end{aligned} \tag{12}$$

which adds constraints to the WF-in-DFT energy expression (Eq. 9) that reflect the nonvariational nature of the approach. The second term, $\sum_m \Lambda_m^{\text{WF,A}} c_m$, contains any constraints, c_m, and their corresponding Lagrange multipliers, $\Lambda_m^{\text{WF,A}}$, that arise from the WF Lagrangian (see for example the MP2 gradient formulation of Ref. 61). The third term places an orthonormality constraint on the DFT MOs, accounting for the incomplete and atom-centered nature of the basis set. The fourth term enforces the Brillouin conditions, $F_{ai} = 0$, associated with self-consistent optimization of the DFT MOs for the full system. Finally, the fifth term enforces the localization condition, $r_{ij} = 0$, that determines the DFT LMOs.

The Lagrangian in Eq. 12 can immediately be used to obtain the derivative of the projection-based WF-in-DFT energy with respect to nuclear positions. (The same formalism can be used to derive responses to electric fields, or to other parameters in the molecular Hamiltonian.) Once the Lagrange multipliers, $\mathbf{z}$, $\mathbf{z}^{\text{loc}}$ and $\mathbf{x}$, have been determined by making $\mathcal{L}$ stationary, the analytical nuclear gradient can be broadly broken down into three components. First, there is the subsystem A WF component which is simply evaluated using existing WF gradient implementations with no modification. No additional terms arise from the DFT method since the embedding potential is held constant in the modified one-electron Hamiltonian (Eq. 10). Second, there is a conventional DFT gradient component, which is calculated using existing implementations with the correct input density matrices. Third, there are multiple embedding components that interface the WF and DFT components. These terms involve solving the coupled-perturbed localization and the coupled-perturbed Kohn-Sham equations, which arise in local correlation[61] and DFT response methods (e.g. TD-DFT) respectively. Evaluation of these terms only require the WF relaxed density, which is already computed when calculating the subsystem A WF nuclear gradient, and the computational complexity of these terms is the same as that of DFT. The full details of the WF-in-DFT analytical nuclear gradient formulation are provided in Ref. 59.

Technical Comments

Choice of Localization Method

Projection-based DFT-in-DFT embedding is formally exact with any disjoint partition of orthogonal orbitals. A multiscale embedding method, in which a spatially localized subsystem is treated at a different level of theory from its surroundings, can be constructed by selecting these subsets from a set of spatially localized orbitals.

Methods based on the Pipek-Mezey criterion,[62] which maintain a chemically intuitive separation between orbitals (e.g. σ-type and π-type), have consistently been found to perform well in projection-based embedding.[29,63] In particular, Knizia's intrinsic bond orbitals (IBOs)[64] have been found to be effective at producing compact LMOs that vary relatively smoothly with respect to nuclear coordinates.[63] We note the quality of the embedding depends on the degree to which the molecular orbitals can be localized; for this reason intrinsically delocalized systems, such as metals, remain a challenge.[57]

Atomic Orbital Basis Set Truncation

As described thus far, projection-based WF-in-DFT embedding reduces the number of occupied LMOs that are correlated at the WF level, but leaves the virtual space untouched. Since the cost of WF methods scales steeply with the number virtual orbitals (e.g. $\mathcal{O}(v^4)$ for CCSD), the spatial locality of subsystem A can be used to reduce the effective size of the virtual space for the embedded WF calculation. One strategy for doing so is to employ local correlation techniques for the WF method;[52,65] local correlation methods include parameters that allow for the control of the lengthscale over which excitations are included.[66] A more general strategy for limiting the size of the virtual space in the WF calculation is to truncate the atomic orbital (AO) basis set employed in the WF-in-DFT calculation.

AO truncation for projection-based embedding have been devised to discard AOs on the basis of either distance from subsystem A[28] or magnitude of contribution to the Mulliken population of subsystem A.[63] The latter method has been found to be particularly simple and robust to employ in practice, determining whether to retain each AO via a single density threshold parameter: if the net Mulliken population – computed using the subsystem A density – of an AO is less than the specified threshold, it is removed from the basis set. In practice, we have found that a density threshold of 1×10^{-4} a.u. provides a good balance between speed and accuracy, but that system-specific sensitivity checks should always be performed.

The energy expression for a projection-based WF-in-DFT calculation with AO truncation is

$$\begin{aligned} &E^{\text{trun}}_{\text{WF-in-DFT}}\left[\tilde{\Psi}^{\text{A,trun}}; \bar{\boldsymbol{\gamma}}^{\text{A,trun}}; \boldsymbol{\gamma}^{\text{A}}, \boldsymbol{\gamma}^{\text{B}}\right] = \\ &\quad E_{\text{WF}}\left[\tilde{\Psi}^{\text{A,trun}}\right] \\ &+ \text{tr}\left[(\tilde{\boldsymbol{\gamma}}^{\text{A,trun}} - \bar{\boldsymbol{\gamma}}^{\text{A,trun}})\mathbf{v}^{\text{trun}}_{\text{emb}}\left[\boldsymbol{\gamma}^{\text{A}}, \boldsymbol{\gamma}^{\text{B}}\right]\right] \\ &- E^{\text{trun}}_{\text{DFT}}\left[\bar{\boldsymbol{\gamma}}^{\text{A,trun}}\right] + E_{\text{DFT}}\left[\boldsymbol{\gamma}^{\text{A}} + \boldsymbol{\gamma}^{\text{B}}\right] \\ &+ \mu \text{tr}\left[\tilde{\boldsymbol{\gamma}}^{\text{A,trun}}\mathbf{P}^{\text{B,trun}}\right], \end{aligned} \tag{13}$$

where terms superscripted by "trun" are represented in the truncated AO basis and those without are evaluated in the full basis. The matrices $\mathbf{v}^{\text{trun}}_{\text{emb}}$ and $\mathbf{P}^{\text{B,trun}}$ are first formed in the full AO basis (Eqs. 2 and 5) and then projected onto the truncated basis by removing the rows and columns that correspond to the truncated basis functions. The matrix $\bar{\boldsymbol{\gamma}}^{\text{A,trun}}$ is the subsystem A one-particle density optimized at the DFT level in the truncated basis. In Eq. 13, the leading order error due to AO truncation is corrected at the DFT-in-DFT level (see Eq. 11 in Ref. 63).

Atomic orbital truncation has been shown to greatly speedup up projection-based WF-in-DFT calculations at a small cost in accuracy in total and relative energies.[58,63,67–69] AO truncation can also be applied to DFT-in-DFT embedding to reduce the cost of the high-level DFT calculation.

Even-Handed Subsystem Partitioning

We now address the practical issue of how best to partition the LMOs between subsystems in

applications to chemical reactions. It is convenient and chemically intuitive to associate a set of atoms with subsystem A, and then to automatically select LMOs corresponding to the subsystem A atoms. Typically, this is done by selecting all LMOs with a significant population on those atoms chosen to comprise subsystem A; populations are typically assigned using the atomic population scheme that corresponds to the localization method used to generate the LMOs (i.e., Mulliken populations for Pipek-Mezey localization[62] and intrinsic atomic orbital populations for IBO localization[64]).

This charge-selection strategy provides a good starting point for determining the subsystem A LMOs, but it becomes problematic when applied to processes for which charge-selected LMOs move into or out of subsystem A as a function of molecular geometry. When this occurs, substantial error is incurred and the projection-based embedded potential energy profile can become discontinuous. Such problems often arise in cases involving bond formation or breaking.

To address this problem, we have recently reported an "even-handed" LMO selection strategy[57] which forms a consensus set of subsystem A LMOs to be used at all geometries along a reaction coordinate. For every geometry, this set contains every LMO that is charge-selected at any geometry. The even-handed LMO selection procedure is automatic, uses information already available at the DFT level, and requires no user input beyond the set of atoms to be embedded (the same input as in the charge-selection method). Even-handed selection has been empirically demonstrated to result in smooth and quantitative energy profiles at the cost of only a few additional LMOs included in subsystem A.[57,58]

Selected Applications

Among the most important aspects of projection-based embedding is that it enables robust and efficient WF-in-DFT calculations in complex chemical systems. To date, projection-based embedding has been used in applications studies from the groups of the authors[58,65,67–71] and others[72–78], including applications to periodic systems.[79,80] We now summarize applications of the method to transition-metal catalysis, enzyme catalysis, and electrochemistry.

Transition-Metal Catalysis

In a first example,[65] WF-in-DFT embedding was employed to investigate a new class of cobalt-based catalysts for hydrogen evolution. A central challenge in the development of inorganic hydrogen-evolution catalysts is to avoid deleterious coupling of the energetics of metal-site reduction from the kinetics of metal-hydride formation.

Collaboration between theory and experiment identified a family of cobalt diimine-dioxime catalysts that shows promise for achieving this aim by introducing an intramolecular proton-shuttle via a pyridyl pendant group (Fig. 2B). For the intramolecular proton-shuttle reaction in this system, WF-in-DFT was found to converge with a high-level subsystem that included only the LMOs on the central transition-metal atoms and its first coordination sphere (Fig. 2C).

For this reaction, Fig. 2A demonstrates the degree to which (local CCSD(T))-in-DFT embedding (red) can remove the qualitative errors of DFT using the B3P86 functional (black), achieving quantitative agreement with local CCSD(T) performed over the full system. Furthermore, the excellent accuracy of the embedding calculation was achieved while reducing the computational cost of the full wavefunction calculation from 20 hours per energy evaluation down to just a single hour per energy evaluation.[65] As is illustrated in this application, transition metal complexes provide very fruitful application domain for projection-based embedding, given that they typically involve subtle electronic structure in the vicinity of the metal that demands a wavefunction theory description, while the surrounding ligand environment is typically both very expensive for wavefunction theories and adequately described using DFT. For these reasons, several other applications of projection-based WF-in-DFT em-

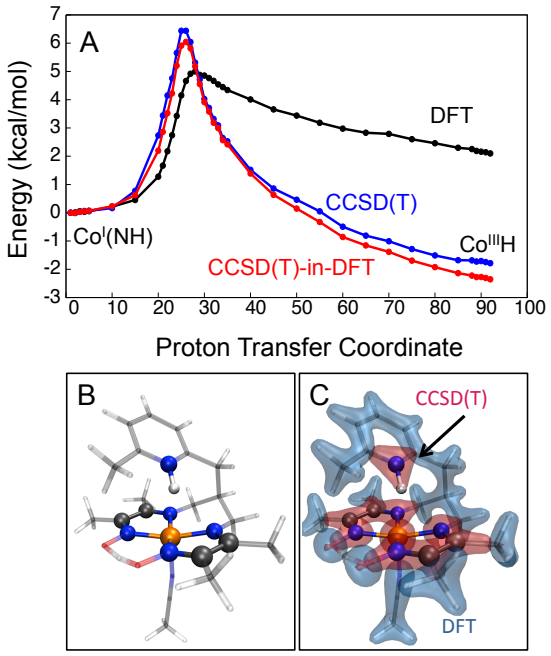

Figure 2: (A) Benchmark energy profiles for the rate-limiting intramolecular proton-transfer reaction in a new class of cobalt diimine-dioxime catalysts, obtained using CCSD(T) (blue), B3P86 (black), and CCSD(T)-in-B3P86 embedding (red). (B) Partitioning of the system in the CCSD(T)-in-B3P86 calculations into atoms that are treated using CCSD(T) (solid ball-stick) and B3P86 (stick). (C) The associated partitioning of the electronic density into subsystems that are treated using CCSD(T) (red) and B3P86 region (blue). Adapted with permission from Ref. 65. Copyright 2016 American Chemical Society.

bedding have also focused on transition-metal complexes.[57,58,70,74]

Enzyme Catalysis

Application of projection-based WF-in-DFT embedding to the reactivity of the citrate synthase enzyme illustrates the potential of the method to help elucidate biochemical mechanisms and predict pharmacological activity (Figure 3).[67,69] DFT is commonly used for the QM region of QM/MM calculations of enzyme systems because of its low cost and the ready availability of implementations. Prediction and understanding of activation barriers using DFT is hampered by the sometimes very large sensitivity on the choice of approximate exchange-correlation functional. A combination of expertise, experience, and careful benchmarking can

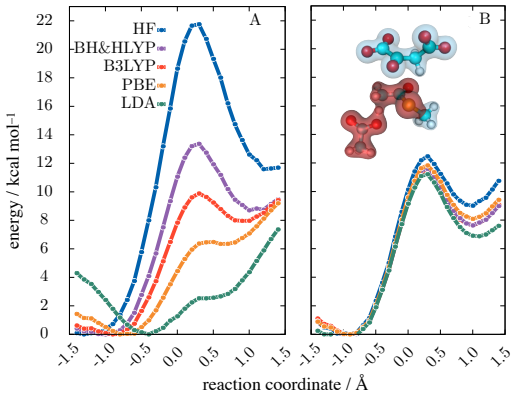

Figure 3: (A) Reaction profiles for proton abstraction from acetyl-coenzyme A in the citrate synthase enzyme, computed using QM/MM with standard DFT approaches (A) and projection-based CCSD(T)-in-DFT embedding (B). The dependence of the predicted activation barrier on choice of functional is almost completely eliminated through projection-based embedding. The QM region is shown as an inset, with the red-shaded density indicating the CCSD(T) region and the remaining environment described using DFT. Note that the CCSD(T)-in-B3LYP and CCSD(T)-in-BH&HLYP curves, red and purple, are indistinguishable. Adapted with permission from Ref. 67. Copyright 2016 American Chemical Society.

help identify a functional that should be reliable for a particular case, but we have found that WF-in-DFT embedding can almost completely eliminate this dependence with modest computational cost.

Specifically, the figure illustrates that whereas DFT with various XC functionals predict qualitatively different reaction energy profiles for the proton abstraction from acetyl-coenzyme A in the citrate synthase (panel A), CCSD(T)-in-DFT embedding provides nearly identical energy curves when the environment is described using DFT with the corresponding XC functionals. It is clear that in this case, the error in the DFT reaction profiles is associated with the local description of the chemical rearrangement,[29] which is robustly corrected using the projection-based WF-in-DFT framework.

Battery Electrolytes

A central challenge in the refinement of lithium-ion batteries is to control cathode-induced oxidative decomposition of electrolyte solvents, such as ethylene carbonate (EC) and dimethyl carbonate (DMC). In recent work, projection-based embedding was used to study the oxidation potentials of neat EC, neat DMC, and 1:1 mixtures of EC and DMC, to overcome qualitative inaccuracies in the electronic densities and ionization energies obtained from conventional KS-DFT methods.[71] The embedding method was implemented as shown in Fig. 4, with a CCSD(T) description of the oxidized molecule, a DFT description of the surrounding molecules, and a molecular-mechanics (MM) description of more distant molecules. Configurations were sampled using classical MD trajectories on the MM force field, and approximately 2000 CCSD(T)-in-DFT-in-MM calculations were performed to obtained the thermal ensemble averages for the oxidation potentials.

It was shown that the ensemble-averaged distributions of vertical IEs are consistent with a linear response interpretation of the statistics of the solvent configurations (Fig. 4C,D), enabling determination of both the intrinsic oxidation potential of the solvents and the corresponding solvent reorganization energies. In-

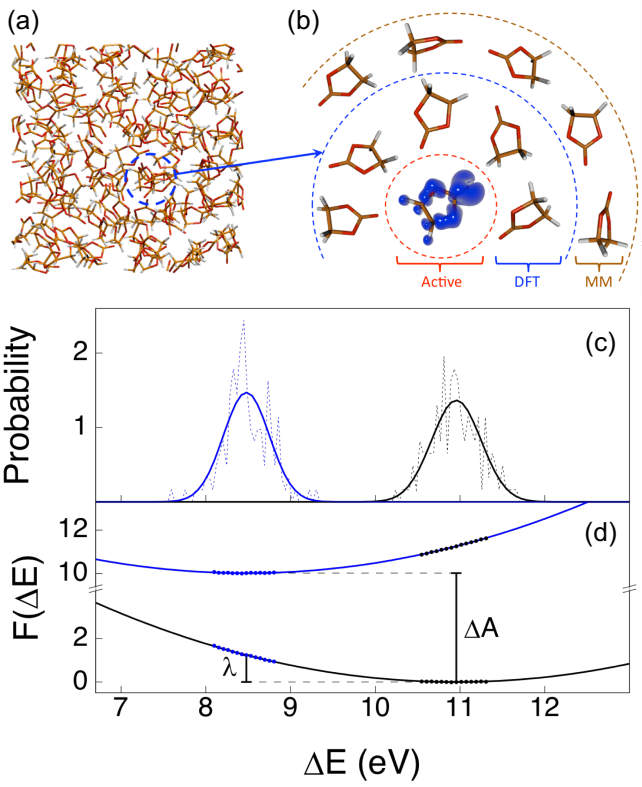

Figure 4: Summary of the embedding protocol in Ref. 71. (a) MD simulations are performed to generate the equilibrium ensemble of solvent configurations. (b) Illustration of the CCSD(T)-in-DFT-in-MM embedding protocol. (c) Equilibrium probability distributions, $P_M(\Delta E)$, of the vertical ionization energies, ΔE, of ethylene carbonyl (EC) molecules, calculated using CCSD(T)-in-B3LYP-in-MM embedding. "M" corresponds either to the reduced EC system (R, black) or the oxidized EC^+ system (O, blue). The distributions have similar standard deviations, demonstrating that the linear response approximation is accurate for this system. The best fit Gaussian distributions, $g_M(\Delta E)$, are indicated in solid lines. (d) Diabatic free energy profiles constructed from the equilibrium distributions shown in (c). The solid lines indicate the results from the Gaussian fits, while the points correspond to simulation data. Adapted with permission from Ref. 71. Copyright 2015 American Chemical Society.

terestingly, it was found that large contributions to the solvation properties of DMC originate from quadrupolar interactions, resulting in a much larger solvent reorganization energy than that predicted using simple dielectric continuum models. Demonstration that the solvation properties of EC and DMC are governed by fundamentally different intermolecular interactions provides useful insight into lithium-ion batteries electrolyte design, with relevance to electrolyte decomposition processes, solid-electrolyte interphase formation, and the local solvation environment of lithium cations.[71]

Outlook and Conclusions

Quantum embedding methods have long been recognized as a promising approach to achieving high-accuracy quantum chemical descriptions while preserving a tractable computational cost. As this Account describes, projection-based WF-in-DFT embedding offers a simple and accurate strategy for reaching this goal that is of practical utility in many chemical applications areas. While many previous studies had recognized that subsystem embedding could be usefully employed via enforcement of subsystem orthogonalization,[48,49] the key advances of the projection-based embedding method[27] were to recognize that the strategy *(i)* could be used to formulate a formally exact method for DFT embedding; *(ii)* could be used to formulate a rigorous and accurate approach to WF-in-DFT embedding; and *(iii)* could be implemented via an extremely simple level-shift projection operator, such that no extra programming is needed to add a new WF method, even at the level of the gradient code.

Continued technical advances, including atomic-orbital basis set truncation[63] and even-handed subsystem partitioning,[57] have improved the efficiency and robustness of the method, and development of the gradient theory further broadens opportunities for chemical applications.[59] The approach is implemented in the widely used Molpro software package,[60] allowing straightforward application of projection-based WF-in-DFT embedding for quantum chemical studies.[58,65,67–80]

Looking forward, we anticipate continued refinement of the projection-based embedding methodology, as well as increasingly widespread application in diverse areas of chemistry, biology, and materials science.

Acknowledgements

S.J.R.L. and M.W. thank the Resnick Sustainability Institute for graduate student and postdoctoral fellowships, respectively. T.F.M. acknowledges support in part from the NSF (Award No. CHE-1611581). F.R.M. acknowledges support in part from the EPSRC (Grant No. EP/M013111/1). T.F.M and F.R.M also thank the DOE for support (Award No. DE-FOA-0001912).

Biographical Information

Sebastian Lee received a B.S. degree in Chemistry and a Minor in Physics from the University of California, Santa Barbara in 2015. He is currently a Ph.D. candidate in Chemistry at the California Institute of Technology in the group of Professor Thomas Miller working on the development and application of DFT-based embedding methodologies. He has received the Caltech CCE Divisional Fellowship, the Resnick Institute Graduate Fellowship and the MolSSI Seed Fellowship.

Matthew Welborn received his Ph.D. from MIT in 2016. He is currently a Resnick Sustainability Institute postdoctoral fellow at Caltech under the supervision of Prof. Thomas Miller. His research interests include embedding and machine learning approaches to the electronic structure problem.

Fred Manby is a Professor of Theoretical Chemistry at the University of Bristol, with research interests in density functional theory, correlated electronic structure theory, and quantum chemistry software development. After a DPhil at the University of York (1997) he worked as a postdoc with Prof Peter Knowles, then at the University of Birmingham. In 2000, Manby was awarded a Royal Society University

Research Fellowship, which facilitated his move to Bristol to establish an independent research career. He has received the Annual Medal of the International Academy of Quantum Molecular Sciences (2007), is a member of the board of the World Association of Theoretical and Computational Chemists (WATOC), and was awarded Marlow and Corday-Morgan Prizes by the Royal Society of Chemistry.

Thomas Miller is a Professor of Chemistry at the California Institute of Technology, with research interests in the development of theoretical and computational methods to study chemical processes that are related to catalysis, battery technologies, and membrane protein biosynthesis. After completing his undergraduate studies at Texas A&M University, he attended graduate school in the UK on a British Marshall Scholarship and received his Ph.D. from Oxford University in 2005. Miller returned to the US for a postdoctoral fellowship at UC Berkeley, joined the Caltech faculty in 2008, and was promoted to full professor in 2013. While at Caltech, he has received awards that include the Sloan Research Fellowship, NSF CAREER Award, Associated Students of Caltech Teaching Award, Dreyfus Teacher-Scholar Award, and the ACS Early-Career Award in Theoretical Chemistry.

References

(1) Curutchet, C.; Mennucci, B. Quantum chemical studies of light harvesting. *Chem. Rev.* **2017**, *117*, 294–343.

(2) Blomberg, M. R. A.; Borowski, T.; Himo, F.; Liao, R.-Z.; Siegbahn, P. E. M. Quantum chemical studies of mechanisms for metalloenzymes. *Chem. Rev.* **2014**, *114*, 3601–3658.

(3) Seh, Z. W.; Kibsgaard, J.; Dickens, C. F.; Chorkendorff, I.; Nørskov, J. K.; Jaramillo, T. F. Combining theory and experiment in electrocatalysis: Insights into materials design. *Science* **2017**, *355*, eaad4998.

(4) Wentzcovitch, R. M.; Stixrude, L. *Theoretical and Computational Methods in Mineral Physics: Geophysical Applications*; Reviews in Mineralogy & Geochemistry; De Gruyter, 2018.

(5) Kroes, G.-J.; Díaz, C. Quantum and classical dynamics of reactive scattering of H_2 from metal surfaces. *Chem. Soc. Rev.* **2016**, *45*, 3658–3700.

(6) Jones, R. O. Density functional theory: Its origins, rise to prominence, and future. *Rev. Mod. Phys.* **2015**, *87*, 897–923.

(7) Warshel, A.; Levitt, M. Theoretical studies of enzymatic reactions: Dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme. *J. Mol. Biol.* **1976**, *103*, 227–249.

(8) Field, M. J.; Bash, P. A.; Karplus, M. A combined quantum mechanical and molecular mechanical potential for molecular dynamics simulations. *J. Comput. Chem.* **1990**, *11*, 700–733.

(9) Vreven, T.; Morokuma, K. In *Chapter 3 Hybrid Methods: ONIOM(QM:MM) and QM/MM*; Spellmeyer, D. C., Ed.; Annu. Rep. Comput. Chem.; Elsevier, 2006; Vol. 2; pp 35 – 51.

(10) Gordon, M. S.; Fedorov, D. G.; Pruitt, S. R.; Slipchenko, L. V. Fragmentation methods: A route to accurate calculations on large systems. *Chem. Rev.* **2012**, *112*, 632–672.

(11) Lambrecht, D. S. Generalizing energy decomposition analysis to response properties to inform expedited predictive models. *Comput. Theor. Chem.* **2019**, *1149*, 24 – 30.

(12) Neuhauser, D.; Baer, R.; Rabani, E. Communication: Embedded fragment stochastic density functional theory. *J. Chem. Phys.* **2014**, *141*, 041102.

(13) Senator, G.; Subbaswamy, K. Density dependence of the dielectric constant of rare-gas crystals. *Phys. Rev. B* **1986**, *34*, 5754–5757.

(14) Cortona, P. Self-consistently determined properties of solids without band-structure calculations. *Phys. Rev. B* **1991**, *44*, 8454–8458.

(15) Wesolowski, T. A.; Warshel, A. Frozen density functional approach for ab-initio calculations of solvated molecules. *J. Phys. Chem.* **1993**, *97*, 8050–8053.

(16) Iannuzzi, M.; Kirchner, B.; Hutter, J. Density functional embedding for molecular systems. *Chem. Phys. Lett.* **2006**, *421*, 16–20.

(17) Jacob, C. R.; Neugebauer, J.; Visscher, L. Software news and update: A flexible implementation of frozen-density embedding for use in multilevel simulations. *J. Comput. Chem.* **2008**, *29*, 1011–1018.

(18) Roncero, O.; de Lara-Castells, M. P.; Villarreal, P.; Flores, F.; Ortega, J.; Paniagua, M.; Aguado, A. An inversion technique for the calculation of embedding potentials. *J. Chem. Phys.* **2008**, *129*, 184104.

(19) Govind, N.; Wang, Y. A.; da Silva, A. J. R.; Carter, E. A. Accurate ab initio energetics of extended systems via explicit correlation embedded in a density functional environment. *Chem. Phys. Lett.* **1998**, *295*, 129–134.

(20) Huang, C.; Carter, E. A. Potential-functional embedding theory for molecules and materials. *J. Chem. Phys.* **2011**, *135*, 194104.

(21) Huang, C.; Pavone, M.; Carter, E. A. Quantum mechanical embedding theory based on a unique embedding potential. *J. Chem. Phys.* **2011**, *134*, 154110.

(22) Elliott, P.; Cohen, M. H.; Wasserman, A.; Burke, K. Density functional partition theory with fractional occupations. *J. Chem. Theory Comput.* **2009**, *5*, 827–833.

(23) Elliott, P.; Burke, K.; Cohen, M. H.; Wasserman, A. Partition density-functional theory. *Phys. Rev. A* **2010**, *82*, 024501.

(24) Goodpaster, J. D.; Ananth, N.; Manby, F. R.; Miller, T. F., III Exact nonadditive kinetic potentials for embedded density functional theory. *J. Chem. Phys.* **2010**, *133*, 084103.

(25) Goodpaster, J. D.; Barnes, T. A.; Miller, T. F., III Embedded density functional theory for covalently bonded and strongly interacting subsystems. *J. Chem. Phys.* **2011**, *134*, 164108.

(26) Goodpaster, J. D.; Barnes, T. A.; Manby, F. R.; Miller, T. F., III Density functional theory embedding for correlated wavefunctions: Improved methods for open-shell systems and transition metal complexes. *J. Chem. Phys.* **2012**, *137*, 224113.

(27) Manby, F. R.; Stella, M.; Goodpaster, J. D.; Miller, T. F., III A simple, exact density-functional theory embedding scheme. *J. Chem. Theory Comput.* **2012**, *8*, 2564–2568.

(28) Barnes, T. A.; Goodpaster, J. D.; Manby, F. R.; Miller, T. F., III Accurate basis-set truncation for wavefunction embedding. *J. Chem. Phys.* **2013**, *139*, 024103.

(29) Goodpaster, J. D.; Barnes, T. A.; Manby, F. R.; Miller, T. F., III Accurate and systematically improvable density functional theory embedding for correlated wavefunctions. *J. Chem. Phys.* **2014**, *140*, 18A507.

(30) Knizia, G.; Chan, G. K.-L. Density matrix embedding: A simple alternative to dynamical mean-field theory. *Phys. Rev. Lett.* **2012**, *109*, 186404.

(31) Knizia, G.; Chan, G. K.-L. Density matrix embedding: A strong-coupling quantum embedding theory. *J. Chem. Theory Comput.* **2013**, *9*, 1428–1432.

(32) Welborn, M.; Tsuchimochi, T.; Van Voorhis, T. Bootstrap embedding: An internally consistent fragment-based method. *J. Chem. Phys.* **2016**, *145*.

(33) Fornace, M. E.; Lee, J.; Miyamoto, K.; Manby, F. R.; Miller, T. F., III Embedded mean-field theory. *J. Chem. Theory Comput.* **2015**, *11*, 568–580.

(34) Lin, H.; Truhlar, D. G. QM/MM: What have we learned, where are we, and where do we go from here? *Theor. Chem. Acc.* **2006**, *117*, 185–199.

(35) Jacob, C. R.; Neugebauer, J. Subsystem density functional theory. *WIREs Comput. Mol. Sci.* **2014**, *4*, 325–362.

(36) Yu, K.; Krauter, C. M.; Dieterich, J. M.; Carter, E. A. *Fragmentation*; Wiley-Blackwell, 2017; Chapter 2, pp 81–117.

(37) Sun, Q.; Chan, G. K.-L. Quantum embedding theories. *Acc. Chem. Res.* **2016**, *49*, 2705–2712.

(38) Wesolowski, T. A. *Computational Chemistry: Reviews of Current Trends*; World Scientific: Singapore, 2006; Vol. 10; pp 1–82.

(39) Götz, A. W.; Beyhan, S. M.; Visscher, L. Performance of kinetic energy functionals for interaction energies in a subsystem formulation of density functional theory. *J. Chem. Theo. Comput* **2009**, *5*, 3161–3174.

(40) Iannuzzi, M.; Kirchner, B.; Hutter, J. Density functional embedding for molecular systems. *Chem. Phys. Lett.* **2006**, *421*, 16–20.

(41) Fux, S.; Jacob, C. R.; Neugebauer, J.; Visscher, L.; Reiher, M. Accurate frozen-density embedding potentials as a first step towards a subsystem description of covalent bonds. *J. Chem. Phys.* **2010**, *132*, 164101.

(42) Nafziger, J.; Wu, Q.; Wasserman, A. Molecular binding energies from partition density functional theory. *J. Chem. Phys.* **2011**, *135*, 234101.

(43) Unsleber, J. P.; Neugebauer, J.; Jacob, C. R. No need for external orthogonality in subsystem density-functional theory. *Phys. Chem. Chem. Phys.* **2016**, *18*, 21001–21009.

(44) Ryabinkin, I. G.; Kananenka, A. A.; Staroverov, V. N. Accurate and efficient approximation to the optimized effective potential for exchange. *Phys. Rev. Lett.* **2013**, *111*, 013001.

(45) Yang, W.; Wu, Q. Direct method for optimized effective potentials in density-functional theory. *Phys. Rev. Lett.* **2002**, *89*, 143002.

(46) Kümmel, S.; Perdew, J. P. Optimized effective potential made simple: Orbital functionals, orbital shifts, and the exact Kohn-Sham exchange potential. *Phys. Rev. B* **2003**, *68*, 035103.

(47) Jensen, D. S.; Wasserman, A. Numerical methods for the inverse problem of density functional theory. *Int. J. Quantum Chem.* **2018**, *118*, e25425.

(48) Phillips, J. C.; Kleinman, L. New method for calculating wave functions in crystals and molecules. *Phys. Rev.* **1959**, *116*, 287–294.

(49) Huzinaga, S.; Cantu, A. A. Theory of separability of many-electron systems. *J. Chem. Phys.* **1971**, *55*, 5543–5549.

(50) Wasserman, A.; Nafziger, J.; Jiang, K.; Kim, M.-C.; Sim, E.; Burke, K. The importance of being inconsistent. *Annu. Rev. Phys. Chem.* **2017**, *68*, 555–581.

(51) Pennifold, R. C. R.; Bennie, S. J.; Miller, T. F., III; Manby, F. R. Correcting

density-driven errors in projection-based embedding. *J. Chem. Phys.* **2017**, *146*, 084113.

(52) Hégely, B.; Nagy, P. R.; Ferenczy, G. G.; Kállay, M. Exact density functional and wave function embedding schemes based on orbital localization. *J. Chem. Phys.* **2016**, *145*, 064107.

(53) Culpitt, T.; Brorsen, K. R.; Hammes-Schiffer, S. Communication: Density functional theory embedding with the orthogonality constrained basis set expansion procedure. *J. Chem. Phys.* **2017**, *146*, 211101.

(54) Henderson, T. M. Embedding wave function theory in density functional theory. *J. Chem. Phys.* **2006**, *125*, 014105.

(55) Khait, Y. G.; Hoffmann, M. R. In *Annu. Rep. Comput. Chem.*; Wheeler, R. A., Ed.; Annu. Rep. Comput. Chem.; Elsevier, 2012; Vol. 8; pp 53 – 70.

(56) Sæther, S.; Kjærgaard, T.; Koch, H.; Høyvik, I.-M. Density-based multilevel Hartree-Fock model. *J. Chem. Theory Comput.* **2017**, *13*, 5282–5290.

(57) Welborn, M.; Manby, F. R.; Miller, T. F., III Even-handed subsystem selection in projection-based embedding. *J. Chem. Phys.* **2018**, *149*, 144101.

(58) Chapovetsky, A.; Welborn, M.; Luna, J. M.; Haiges, R.; Miller, T. F., III; Marinescu, S. C. Pendant hydrogen-bond donors in cobalt catalysts independently enhance CO_2 reduction. *ACS Cent. Sci.* **2018**, *4*, 397–404.

(59) Lee, S. J. R.; Ding, F.; Manby, F. R.; Miller, T. F., III Analytical Gradients for Projection-Based Wavefunction-in-DFT Embedding. *arXiv e-prints* **2019**, arXiv:1903.05830.

(60) Werner, H.-J.; Knowles, P. J.; Knizia, G.; Manby, F. R.; Schütz, M.; Celani, P.; Györffy, W.; Kats, D.; Korona, T.; Lindh, R.; Mitrushenkov, A.; Rauhut, G.; Shamasundar, K. R.; Adler, T. B.; Amos, R. D.; Bennie, S. J.; Bernhardsson, A.; Berning, A.; Cooper, D. L.; Deegan, M. J. O.; Dobbyn, A. J.; Eckert, F.; Goll, E.; Hampel, C.; Hesselmann, A.; Hetzer, G.; Hrenar, T.; Jansen, G.; Köppl, C.; Lee, S. J. R.; Liu, Y.; Lloyd, A. W.; Ma, Q.; Mata, R. A.; May, A. J.; McNicholas, S. J.; Meyer, W.; Miller, T. F., III; Mura, M. E.; Nicklass, A.; O'Neill, D. P.; Palmieri, P.; Peng, D.; Pflüger, K.; Pitzer, R.; Reiher, M.; Shiozaki, T.; Stoll, H.; Stone, A. J.; Tarroni, R.; Thorsteinsson, T.; Wang, M.; Welborn, M. MOLPRO, version 2018.2, a package of ab initio programs. 2018; see http://www.molpro.net.

(61) Schütz, M.; Werner, H.-J.; Lindh, R.; Manby, F. R. Analytical energy gradients for local second-order Møller-Plesset perturbation theory using density fitting approximations. *J. Chem. Phys.* **2004**, *121*, 737–750.

(62) Pipek, J.; Mezey, P. G. A fast intrinsic localization procedure applicable for ab initio and semiempirical linear combination of atomic orbital wave functions. *J. Chem. Phys.* **1989**, *90*, 4916–4926.

(63) Bennie, S. J.; Stella, M.; Miller, T. F., III; Manby, F. R. Accelerating wavefunction in density-functional-theory embedding by truncating the active basis set. *J. Chem. Phys.* **2015**, *143*, 024105.

(64) Knizia, G. Intrinsic Atomic Orbitals: An unbiased bridge between quantum theory and chemical concepts. *J. Chem. Theory Comput.* **2013**, *9*, 4834–4843.

(65) Huo, P.; Uyeda, C.; Goodpaster, J. D.; Peters, J. C.; Miller, T. F., III Breaking the correlation between energy costs and kinetic barriers in hydrogen evolution via a cobalt (pyridine-diimine-dioxime) catalyst. *ACS Catal.* **2016**, *6*, 6114–6123.

(66) Schütz, M.; Hetzer, G.; Werner, H.-J. Low-order scaling local electron correlation methods. I. Linear scaling local MP2. *J. Chem. Phys.* **1999**, *111*, 5691–5705.

(67) Bennie, S. J.; van der Kamp, M. W.; Pennifold, R. C.; Stella, M.; Manby, F. R.; Mulholland, A. J. A projector embedding approach for multiscale coupled-cluster calculations applied to citrate synthase. *J. Chem. Theory Comput.* **2016**, *12*, 2689–2697.

(68) Bennie, S. J.; Curchod, B. F.; Manby, F. R.; Glowacki, D. R. Pushing the limits of EOM-CCSD with projector-based embedding for excitation energies. *J. Phys. Chem. Letts.* **2017**, *8*, 5559–5565.

(69) Zhang, X.; Bennie, S. J.; van der Kamp, M. W.; Glowacki, D. R.; Manby, F. R.; Mulholland, A. J. Multiscale analysis of enantioselectivity in enzyme-catalysed 'lethal synthesis' using projector-based embedding. *Roy. Soc. Open Sci.* **2018**, *5*, 171390.

(70) Stella, M.; Bennie, S. J.; Manby, F. R. Computational study of adsorption of cobalt on benzene and coronene. *Mol. Phys.* **2015**, *113*, 1858–1864.

(71) Barnes, T. A.; Kaminski, J.; Borodin, O.; Miller, T. F., III Ab initio characterization of the electrochemical stability and solvation properties of condensed-phase ethylene carbonate and dimethyl carbonate mixtures. *J. Phys. Chem. C* **2015**, *119*, 3865–3880.

(72) Parrish, R. M.; Gonthier, J. F.; Corminbœuf, C.; Sherrill, C. D. Communication: Practical intramolecular symmetry adapted perturbation theory via Hartree-Fock embedding. *J. Chem. Phys.* **2015**, *143*, 051103.

(73) Meitei, O. R.; Heßelmann, A. Intramolecular interactions in sterically crowded hydrocarbon molecules. *J. Comput. Chem.* **2017**, *38*, 2500–2508.

(74) de Lima Batista, A. P.; de Oliveira-Filho, A. G. S.; Galembeck, S. E. Photophysical properties and the NO photorelease mechanism of a ruthenium nitrosyl model complex investigated using the CASSCF-in-DFT embedding approach. *Phys. Chem. Chem. Phys.* **2017**, *19*, 13860–13867.

(75) Yao, K.; Herr, J. E.; Brown, S. N.; Parkhill, J. Intrinsic bond energies from a bonds-in-molecules neural network. *J. Phys. Chem. Lett.* **2017**, *8*, 2689–2694.

(76) Chen, X.; Goldsmith, C. F. Predictive kinetics for the thermal decomposition of RDX. *Proceedings of the Combustion Institute* **2019**, *37*, 3167 – 3173.

(77) Lin, L.; Zepeda-Núñez, L. Projection based embedding theory for solving Kohn-Sham density functional theory. *arXiv e-prints* **2018**, arXiv:1807.08859.

(78) Böckers, M.; Neugebauer, J. Excitation energies of embedded open-shell systems: Unrestricted frozen-density-embedding time-dependent density-functional theory. *J. Chem. Phys.* **2018**, *149*, 074102.

(79) Libisch, F.; Marsman, M.; Burgdörfer, J.; Kresse, G. Embedding for bulk systems using localized atomic orbitals. *J. Chem. Phys.* **2017**, *147*, 034110.

(80) Chulhai, D. V.; Goodpaster, J. D. Projection-based correlated wave function in density functional theory embedding for periodic systems. *J. Chem. Theory Comput.* **2018**, *14*, 1928–1942.